BMC Bioinformatics

RESEARCH

Open Access

Degeneracy measures in biologically plausible random Boolean networks

Basak Kocaoglu[1,3*] and William H. Alexander[1,2,3]

*Correspondence:
bkocaoglu2018@fau.edu
[1] Center for Complex Systems and Brain Sciences, Florida Atlantic University, Boca Raton, FL, USA
Full list of author information is available at the end of the article

Abstract

Background: Degeneracy—the ability of structurally different elements to perform similar functions—is a property of many biological systems. Highly degenerate systems show resilience to perturbations and damage because the system can compensate for compromised function due to reconfiguration of the underlying network dynamics. Degeneracy thus suggests how biological systems can thrive despite changes to internal and external demands. Although degeneracy is a feature of network topologies and seems to be implicated in a wide variety of biological processes, research on degeneracy in biological networks is mostly limited to weighted networks. In this study, we test an information theoretic definition of degeneracy on random Boolean networks, frequently used to model gene regulatory networks. Random Boolean networks are discrete dynamical systems with binary connectivity and thus, these networks are well-suited for tracing information flow and the causal effects. By generating networks with random binary wiring diagrams, we test the effects of systematic lesioning of connections and perturbations of the network nodes on the degeneracy measure.

Results: Our analysis shows that degeneracy, on average, is the highest in networks in which ~ 20% of the connections are lesioned while 50% of the nodes are perturbed. Moreover, our results for the networks with no lesions and the fully-lesioned networks are comparable to the degeneracy measures from weighted networks, thus we show that the degeneracy measure is applicable to different networks.

Conclusions: Such a generalized applicability implies that degeneracy measures may be a useful tool for investigating a wide range of biological networks and, therefore, can be used to make predictions about the variety of systems' ability to recover function.

Introduction

Biological systems can adjust their functioning dynamically in face of changing circumstances. However, such functional adjustments are constrained by the structural properties of the components that perform the functions [1–7] as well as the topology of the system. Biological systems as complex networks have evolved multiple strategies to achieve a 'working' reconfiguration of the components that promotes survival through shifts in environmental contingencies [8–15]. One strategy is redundancy which means

that a system has multiple structurally identical components serving the same function [16–19]. Systems may also utilize multifunctionality (or alternatively, pluripotentiality [20, 21]) that is the capacity for a single component to serve multiple functions [22–26]. Another strategy for biological systems to respond flexibly to perturbations is called degeneracy [9, 10, 27–29].

Degeneracy (or alternatively, 'distributed redundancy' [30], 'distributed robustness' [31], 'functional redundancy' [25], 'extrinsic buffering' [32]) describes the ability of components in a biological system that are structurally different to carry out the same or similar functions [20, 21, 21, 28, 33–42]. As difference in the structure implies different functions [43], under certain conditions degenerate components do not necessarily show functional variety [44, 45] but instead, each degenerate component is responsible for a (set of) function(s) which is initially determined by their biochemical(/physical) structure [9, 21, 31]. Unlike multifunctionality and redundancy, degeneracy implies a change in the role assignments among the components such that the system can continue to function even when its normal processes have been compromised.

A biological network with high degeneracy means that the system can show the same macroscopic behavior following a lesion even though the underlying network dynamics are significantly different. In other words, if the system is highly degenerate, after a lesion, the function can be recovered by a structurally different (i.e., performing a different function under normal conditions) component taking over a new function. For example, in the brain many different neural clusters can affect the same motor outputs, and if some of the brain areas are damaged, an alternative (non-redundant) pathway can be recruited in order to generate functionally equivalent behaviors [46–50]. Degeneracy thus suggests how biological systems can thrive despite changes to internal and external demands.

It has been shown that degeneracy also plays a role in complexity and evolvability of biological systems [8, 10, 27, 28, 32, 36, 51]. Higher levels of degeneracy correlate with an increase in the degree of both the functional integration and local segregation of a system, and therefore, higher degeneracy is accompanied by higher degree of complexity of the systems [33].While local segregation (namely, functional specialization) enables system to be flexible against environmental stress (due to diversity of functions), functional integration allows system to be robust [52–54]. If a component, or a group of components are compromised in a highly degenerate system, functions can be reassigned among distinct elements (that are locally segregated) while the macro-level behavior (which requires the system to be functionally integrated) is conserved. This adaptability brings an obvious advantage over the course of natural selection [46].

To measure degeneracy in systems, Tononi et al. [33] introduced a quantitative measure for neural networks (see also alternatives [55, 56]) using an information theoretic approach. Information theory [57–59] provides a set of tools to describe how information is processed in systems. It allows us to measure the statistical (in)dependencies in terms of the information content of the components. Mutual information (MI), which is a measure provided by information theory, can also capture nonlinear dependencies(/relationships) that are not detectable by correlation analysis [57, 60]. However, direction of the interaction between the components cannot be discerned from MI alone [33, 60, 61]. Incorporating MI with systematic perturbations (to determine directionality),

degeneracy is formalized [33] in terms of the causal effects of the changes in the state of the subsets (components and/or subgroups of components) on the system's output. If the output activity of the system is not affected by the change (e.g., perturbation) in a subset's state, then the system is highly degenerate with regard to the function that is performed by that subset. This information theoretic measure of degeneracy is, first, applied to highly abstract networks in the work by Tononi et al. [33] which is followed by applications to the weighted networks with a high degree of biological fidelity (e.g., Hodgkin–Huxley type neural networks [40] and genetic networks with epistasis [62]).

Although it has been shown that degeneracy as a network property exists at different levels of biological organization (from molecules to behavior [30, 36, 46]), a quantitative analysis of degeneracy at such levels is sparse and methods are individualized to specific cases (see the different versions of degeneracy measurements in other works [10, 55, 63]). In systems biology, information theoretic measures are widely applied to many problems [64], yet, to date, there is not a comprehensive study applying this measure for biologically realistic networks other than networks with weighted connections.

Neural networks offer one example of how biological systems can incorporate degeneracy to ensure survival after being damaged. However, other biological networks are likewise capable of recovering partial or full function following damage. For example, in between-species interaction networks a species loss can be compensated by other species contributing to ecosystem functioning [65]. Likewise, on a smaller scale, it has been shown that loss of functioning in some (non-redundant) genes has a weak or no effect on the fitness of the gene networks [8]. Although degeneracy might not be detectable under normal conditions, after perturbating or lesioning the biological networks, changes in the environment may also evoke degenerate responses ('degeneracy lifting' effect [9]). Environmental (evolutionary) pressure in receptor-signal transduction networks [55] can push the signaling pathways to reconfigure into a degenerate form.

Although degeneracy is a feature of biological gene networks, it is unclear whether models of gene networks can be analyzed using the same information theoretic approaches as used for neural network models. A GRN is a network of gene–gene interactions through their regulators that control the gene expression levels of the products (mRNAs and proteins) which, ultimately, determine the cell fate (final cell type, i.e., function of the cell) [66]. Measurements of degeneracy at the level of gene transcription control may provide insights on how functions of genes that determine the cell function, can be recovered as a consequence of the network properties (GRN topology).

Random Boolean network (RBN) models, as discrete models of GRNs, are well-suited to study degeneracy since with RBNs we can induce and trace the effects of targeted lesions while environmental/external pressure is a parameter that can be controlled over in silico experiments. Unlike neuronal networks where edges are (synaptic) weight vectors, RBNs have a static wiring diagrams [67] governed by logic equations that represent the functions of gene regulatory factors (e.g., transcription factors). Logic equations describe the underlying network architecture. For example, for a simple network of 3 genes, if gene G1 is regulated both directly by G3 and through an indirect link from G2, this architecture is represented by the logic function of "G1 = (G2 OR G3)". Likewise, if there is an inhibitory regulation of G1 through G2 while the same architecture is preserved from the network described above, this structure can be represented by the logic

equation of "G1 = ((NOT G2) OR G3)". Since each state of gene expression is the direct outcome of the activity of (regulatory interactions in) the previous state, one can assess the effects of circuit architecture on gene expression levels [66]. Hence, in RBNs, it is feasible to trace the information flow at each (discrete) time step and so, causal influences.

While RBNs have been widely used in system biology to study GRNs, recent work [68] has sought to match the complexity of gene expression patterns that are observed in developmental processes (such as, mammalian cortical area development [69]) by incorporating additional biological details to simulations. This is attained by converting Boolean models to equivalent systems of differential equations that describe synergistic relations in gene regulatory processes [68]. Unlike widely-used in-silico single-cell gene expression data generators such as GeneNetWeaver [70] where simulated networks are re-constructed from a limited number of known structures (from *E. coli* and *S. cerevisiae* [70]), here, the network structure is defined by a (Boolean) ruleset that can be generated and manipulated by users [68].

In this study we test to what extent the information theoretic measure of degeneracy applies to RBNs as well as continuous models derived from RBN approaches. Furthermore, we test systematic lesions in randomly generated Boolean networks while varying the number of perturbed nodes. This enables us to explore how degeneracy quantitively changes as a function of interventions to the nodes and induced topological alterations in the networks. Results show that degeneracy measures, in most cases, can be applied to networks based on Boolean logic in addition to more typical weighted networks.

Results

Lesioning

We first investigate the effects of systematic lesioning on measures of degeneracy in randomly wired RBNs. This is consistent with classical models of RBNs [67] where the system is deterministic and no noise term is involved. Edges between network nodes were lesioned in two ways. In type-1 lesioning, only incoming edges were lesioned incrementally while it is possible (due to the pseudo-random algorithm to generate the logic functions) that outgoing edges stayed intact (for details see "Methods"). In the second type of lesioning (hence the name type-2 lesion), we lesioned all incoming and outgoing edges for randomly chosen nodes incrementally.

For classical RBNs, both types of lesioning result in an increase on degeneracy measures (Fig. 1). That is, removing edges from the network paradoxically increased degeneracy. One reason for this may be that progressive lesioning increases the chances of a node becoming isolated. In the context of classical RBNs, the activity of the isolated node remains constant over time. However, a constant activity profile may also be a result of a function of a (set of) node(s) that is highly robust and yields same output over time. Therefore, the degeneracy metric might fall short in cases where the outputs of the isolated nodes remain static over time. Here, we hypothesize that degeneracy only works as a measure when all units have varying activity.

To test this hypothesis, we induced a small probability that the activity each node in the network would flip from 1 to 0, or 0 to 1, on each iteration (see "Methods" for details). Results show that the networks with type-1 lesions decrease in average degeneracy values as the cut percentage increases (Fig. 2). Similarly, for type-2 lesions, average

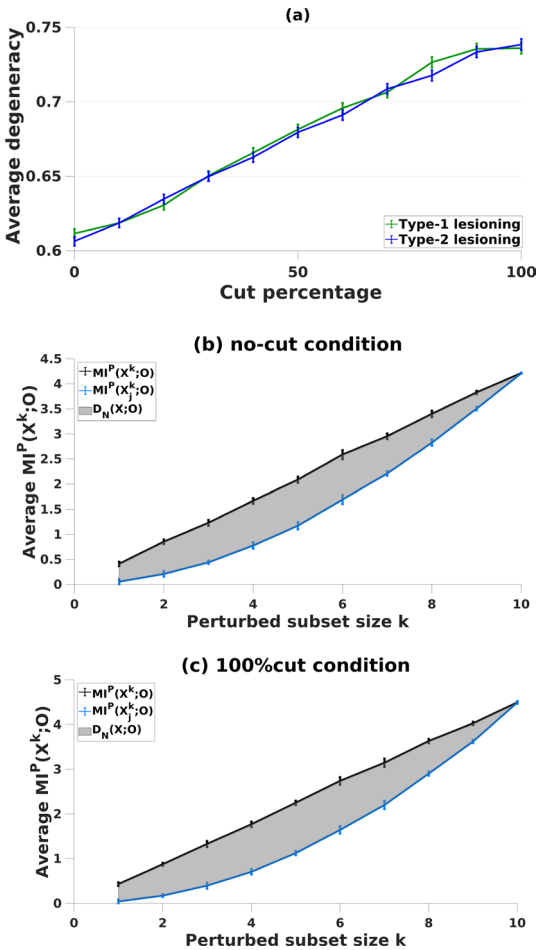

Fig. 1 Average degeneracy for type-1 (green line) and type-2 (blue line) lesioning in classical RBNs (**a**). Degeneracy, (grey area) is computed as the average MI between subsets of X and O under perturbation over increasing perturbed subset size k, in networks with **b** no lesions and with **c** 100%-cut condition. In panel **a**, on the x axis, the cut percentage represents the affected number of nodes (in total of 10 nodes) whose edges are lesioned given the networks. As cuts becomes larger, degeneracy increases. In panels **b** and **c**, degeneracy is calculated according to definition given in "Methods" Eq. 3

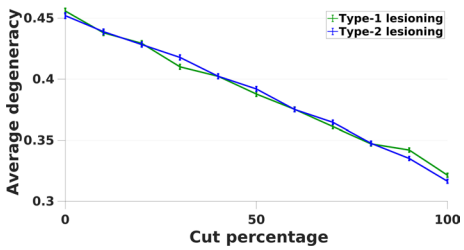

Fig. 2 Average degeneracy values compared between type-1 (green line) and type-2 (blue line) lesioning in RBNs with varying unit activity. On the x axis, the cut percentage represents the affected number of nodes (in total of 10 nodes) whose edges are lesioned given the networks. For both lesioning types, degeneracy was lowest in the 100% cut condition where edges of all nodes (10) were cut

degeneracy decreases as a function of lesioning. This validates that degeneracy emerges as a network property for RBNs, but only when the activity of all nodes in the network changes over time.

Furthermore, we hypothesized that there could be a difference between the effects of lesioning types as a direct consequence of the partial lesioning (the outgoing edges are preserved) in type-1 condition which can lead some nodes to become dead ends since the activity ends in those nodes. Active nodes without incoming edges means that such nodes do not serve a function, and this eventually would result in lower degeneracy values. The comparison of two lesioning types, in Fig. 2, demonstrates that these lesioning types are not significantly different in their effects on degeneracy ($f(1,10)=0$, $p>0.05$).

Perturbation

Degeneracy is calculated as the area between the average $MI^P(X^k;O)$ (mutual information, MI, between the portion of entropy shared by the system for *each* perturbed subset k and the output O) and overall-MI (mutual information between the system and its output) for different perturbed subset sizes, k (see Eq. 3 in "Methods"). This area shows a characteristic shape of the degeneracy function: a non-zero value that declines to zero as perturbed subset size k approaches $k=O$, following an increase that is "higher than would be expected from a linear increase" [33]. This characteristic shape has furthermore been replicated in networks composed of in Hodgkin–Huxley neurons [40].

In weighted networks with no connections, the average (overall-) MI shows a linear increase where degeneracy is zero [33]. Here, this condition (i.e., 100%-of-edges-cut) is compared for RBNs with varying unit activity. Our results for no-edges-cut condition and the characteristic profile of degeneracy are comparable to corresponding findings in previous studies (Fig. 3).

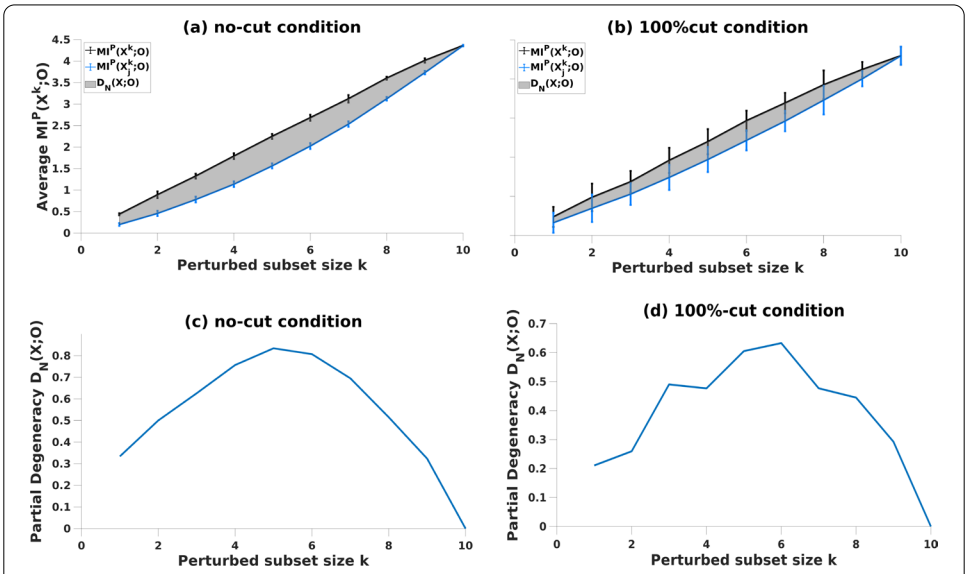

Fig. 3 Degeneracy, (grey area) is computed (see Eq. 3 in "Methods") as the average MI between subsets of X and O under perturbation over increasing perturbed subset size k, in networks with **a**, **c** no lesions and with **b**, **d** 100%-cut condition

Further inspection of partial degeneracy (see "Methods" for details) values from individual simulations showed that partial degeneracy can have a negative value in some instances (only two such conditions captured here for comparison, Fig. 4a, b). Although we have observed that partial degeneracy can be negative for different cut-conditions and different sizes of perturbed subset, when MI is averaged over the simulations for all the perturbed subset sizes k, $\langle MI^P(X^k;O)\rangle$, degeneracy $D_N(X; O)$ was above zero in all conditions.

Interactions of lesions and perturbations

Increases in both lesion extent and the number of nodes perturbed give rise to different metrics in terms of degeneracy, raising the question of how these two factors may interact. We therefore conducted additional simulations in which each lesion condition (0–100%, see "Methods") was crossed with each perturbation condition (k = 1–10). Figure 5a, b show how average partial degeneracy changes as a function of perturbation subset size k, given cut percentages. For both lesioning types, partial degeneracy peaks around when half of the nodes ($k \sim 5$) are perturbed in the system. The measure of degeneracy can detect existing isofunctionality between the different structures only when one of the structures is perturbed. When half of the nodes in the system are perturbed, we, thereby, maximize the probability of selecting/measuring the right structure for given degeneracy in cases where network structure is random or unknown.

Results for continuous models with RBN connectivity scheme (cBNs)

We apply the same perturbation and lesioning procedures to RBNs that generate continuous expression data. These networks also have initial variance due to the model (see

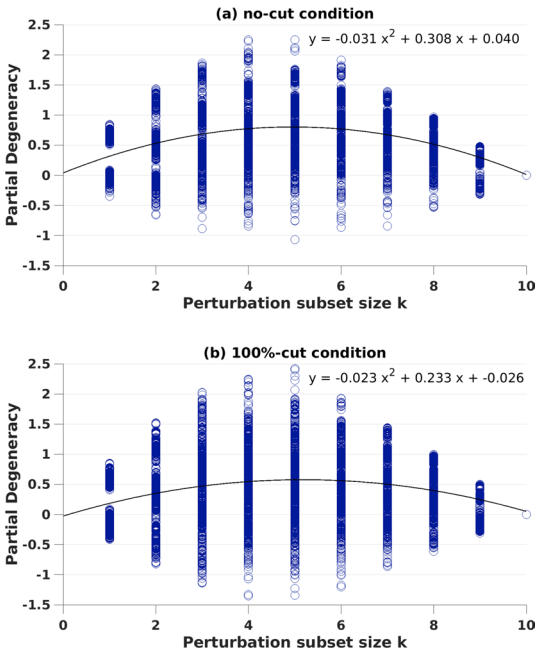

Fig. 4 Partial degeneracy values from individual networks for each perturbed subset size k. Data from 10 (k) × 1000 simulations of networks with **a** no-cut condition and **b** 100%-cut condition

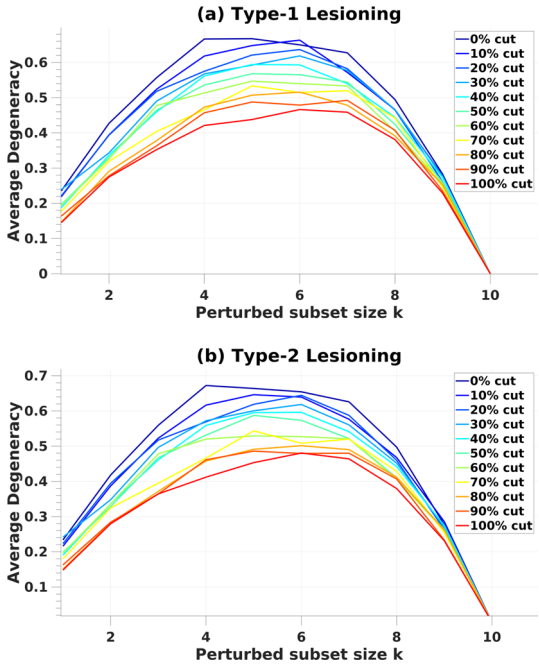

Fig. 5 Average degeneracy computed as a function of perturbation subset size k in type-1 lesioning (**a**) and type-2 lesioning (**b**). Each line represents the cut condition for lesioned edges given the percentage of number of nodes in networks

details in "Methods" and Additional file 1) which is a system of stochastic differential equations (SDEs).

The comparison of the two lesioning types in cBNs (Fig. 6a), demonstrates that both conditions have similar effects on the average degeneracy, where there is a no significant difference (two-way analysis of variance, ANOVA) found between both types of lesioning ($f(1,10) = 9.49$, $p > 0.01$). However, average degeneracy varies with cut conditions.

The degeneracy measures—the area between the average $MI^P(X^k;O)$ *and overall-MI*—demonstrate similar characteristics (Fig. 6c, d), and thus are comparable, to RBN measures. For the 100%-cut condition (Fig. 6c), degeneracy is closest to zero, consistent with findings in previous studies [33]. Likewise, partial degeneracy shows a negative value in some instances (Fig. 6f, g) although on average (MI) degeneracy was above zero for all conditions and subset sizes k. Moreover, comparable to RBNs, the peak for partial degeneracy measures is around when half of the nodes ($k \sim 5$) are perturbed in the system for both lesioning types (Fig. 6).

Discussion

Although a variety of types of biological networks are thought to exhibit degeneracy, previous theoretical work has primarily focused on networks with weighted connections [33, 40, 62]. In this study, we demonstrate that degeneracy measures are also suitable for RBNs, with some caveats. Although our simulations largely replicated previous studies investigating degeneracy in neurally-inspired networks, RBNs use Boolean logic operators rather than weighted connections to determine function. It therefore might have been the case that information-theoretic approaches developed

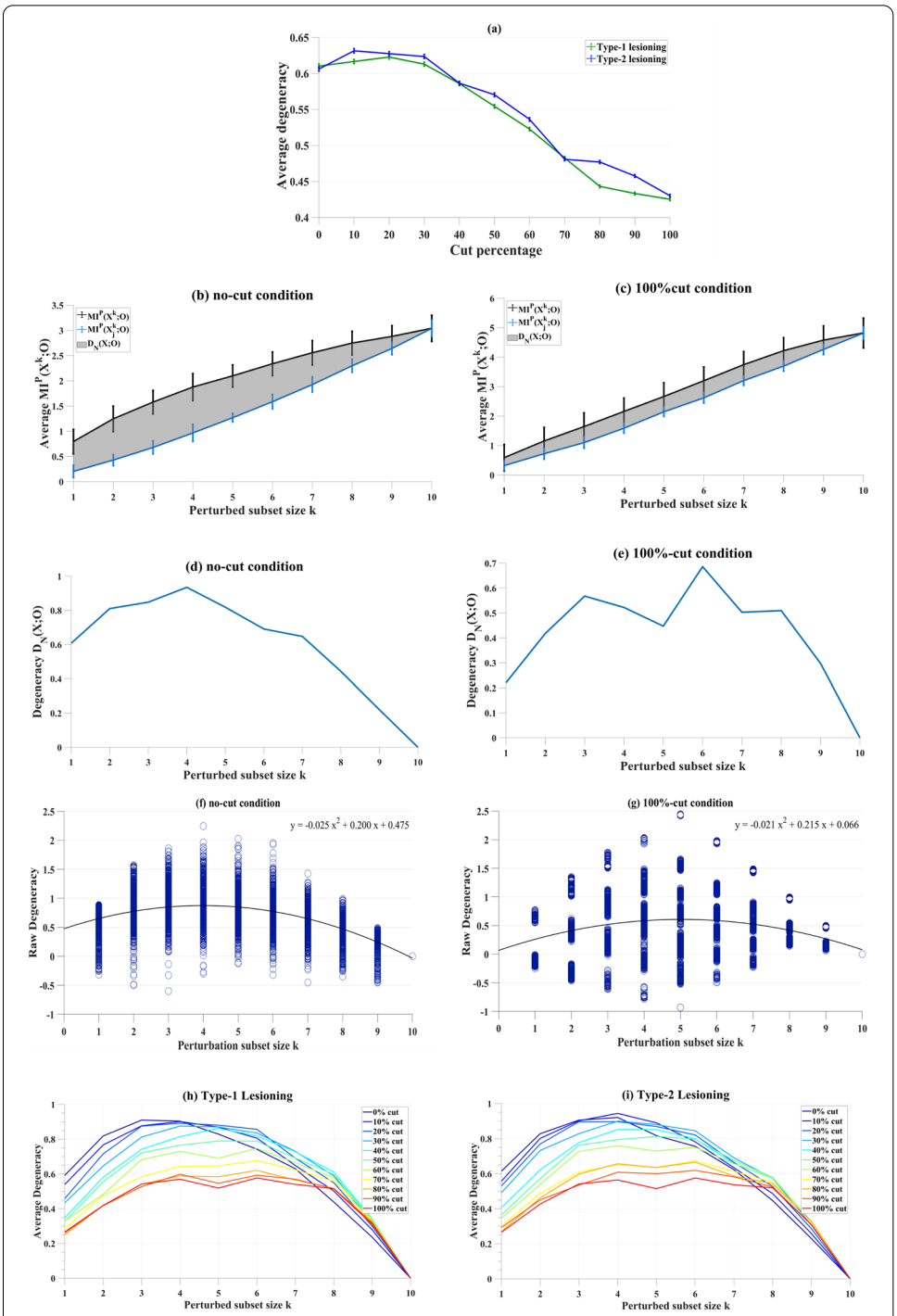

Fig. 6 Panel for cBNs with type-1 and type-2 lesioning (**b**). In **a**, average degeneracy values compared between type-1 (green line) and type-2 (blue line) lesioning. Degeneracy, (grey area) is computed (see Eq. 3 in "Methods") as the average MI between subsets of *X* and *O* under perturbation over increasing perturbed subset size *k*, in networks with **b**, **d** no lesions and with **c**, **e** 100%-cut condition. Panels **f**, **g** show partial degeneracy values from individual networks for each perturbed subset size k. The distribution of partial degeneracy values for **g** shows clear modes in the data where there is an overlap of the perturbed subset and output sheet. Average degeneracy computed as a function of perturbation subset size k in type-1 lesioning (**h**) and type-2 lesioning (**i**)

for one class of networks might not have generalized correctly to a new class. By replicating previous findings using RBN (which are discrete systems) and cBNs, we demonstrate that information-theoretic approaches are applicable to a broad range of network types.

Network and graph theoretic approaches, frequently formulated in terms of information theory, have been applied extensively to neuroscience [52–54, 61, 71–80] to predict individual differences, consequences of lesions, and ability to recover function following injury. Extending this approach to the study of GRNs opens the door to investigating the consequences of, and possible remedies for, genetic dysfunction. Because most genetic functions are performed by subsets of many components within functional modules [5, 81], diseases may emerge due to disorganization of the components in these modules. Degeneracy measures can be recruited for predicting and inducing topological modifications (for example, 'rewiring of diseased modules' [81]) to achieve desired functional outcomes that have clinical significance, such as enhanced pharmaceuticals with better drug targets.

In addition to replicating previous results, we explored the impact of systematically manipulating network connectivity (lesioning) while decomposing degeneracy by size of the perturbed subset. In doing so, we identify a potential interaction between the number of perturbed nodes and the magnitude of the impact of lesions on partial network degeneracy. In networks in which ~ 20% of the connections are lesioned while 50% of the nodes are perturbed, it is observed that average partial degeneracy reaches its highest value among all other cut conditions and for all perturbed subset sizes. These are likely conditions in which sufficient connectivity (here, $k = 5$, ~ 10%-cut for RBNs and ~ 20%-cut for cBNs) allows for the expression of degenerate structures without compromising their function, whereas more lesioning would diminish both primary and degenerate structures, and more perturbation would confound the functions of the nodes. Likewise, if perturbation (of the number of nodes) is smaller, not all possible degenerate structures might be expressed in the network or, when the lesioning is less, degenerate structures might be unobservable since most of the primary structures are intact.

In the literature it has been shown that additional damage (gene/node deletions) can restore the function of previously compromised (metabolic) networks [82–84]. Here, we show that progressively lesioning a network results in reduced partial degeneracy across all perturbation subsets. However, when a perturbation subset includes approximately half of the nodes in the network, we observe an increase in partial degeneracy relative to larger or smaller perturbation subsets. This finding suggests that it might be possible to determine an optimum degree of lesioning and perturbation given a network to achieve higher degeneracy in the systems. Thus, partial degeneracy measures might be helpful to develop strategies to predict how to recover the function after damage.

As originally conceived, degeneracy was intended to capture the idea that identical functions could be carried out by distinct network structures. Intuitively, therefore, degeneracy would seem to have a lower bound at zero—in a network with no degeneracy, all structures would serve their own individual functions, and perturbation of those structures would disrupt network output related to the function served.

Although on average degeneracy in our simulations tended to be above or equal to zero, we observed individual simulations in which partial degeneracy values were below zero. In previous studies [40, 62], negative degeneracy has been observed especially for network models with increased biological fidelity.

One possible reason for the observation of negative degeneracy may be that the information-theoretic measure for degeneracy was originally developed for and tested on neural networks with no initial variance. As the biological fidelity of the models (thus, inherent variance in the systems) increases, for some conditions (lower coupling and lower connection probability [40] and networks with lower complexity [62]) negative degeneracy has been shown. However, we have not observed such effects on overall degeneracy measurements where 1000 MI values for each possible subset size k were averaged across random networks. Mathematically, degeneracy gets a negative value when the portion (k/n) of the MI between the whole system (n) and the output sheet ($n/2$) is higher than the average of the MI between the (perturbed) subset of the system (k) and the output sheet ($n/2$). However, the biological meaning/equivalence of negative degeneracy remains unclear. For studying more biologically realistic complex networks, adjustments in the tools for quantifying degeneracy may be needed.

Another challenge for degeneracy metrics is that, when the units have no varying activity over time due to complete isolation from rest of the network, this may result in higher partial degeneracy values. In that case, the degeneracy measure cannot accommodate the difference between robustness of a node's output and time-invariant output of an isolated node. To account for such an effect, the 'function outcome' definition in degeneracy tools may need to be revised.

Furthermore, in this study, the total number of nodes given a network was set to 10 which is comparable to the network sizes studied in previous works on degeneracy [33, 40, 62]. However real systems such as biological networks often consist of thousands of connected nodes. When we scale up our network simulations to networks with 100 nodes (while other parameters remained the same), we observed that the entropy calculated for conditions in which the number of perturbed units is greater than around 20 reaches an upper bound (see Additional file 1: Text 4 for simulation results).

Since calculating degeneracy depends on accurately calculating entropy over all possible perturbed subset sizes, the solution seems to be to increase the number of timesteps in our simulations. However, we note that when the number of perturbed units is, for example, 100, there are 2^{100} ($\sim 10^{29}$) unique states, indicating the absolute minimum number of timesteps needed to visit each network state once (and accurately estimating entropy would likely require simulations of length 10^{30} timesteps). Simply put, it is unfeasible to calculate degeneracy for even moderately sized networks in the way described by previous authors. We acknowledge this as a weakness of the degeneracy measures that applies both to RBNs as well as weighted network, and suggest that future work should aim at developing definitions of degeneracy that can be applied to larger networks.

Methods

Network architecture

RBNs were initially proposed as simplified models for gene regulatory networks by Kauffmann [85] where network nodes represent the genes, and the edges represent the

regulatory functions. A RBN constitutes a discrete dynamical system that has N nodes, each with K incoming edges (hence, also referred to as N-K models). Each node (gene) can be ON or OFF (1 or 0); a network of N binary nodes therefore has 2^N distinct states [85]. This system is state determined [85] according to Boolean functions that are assigned to each node randomly (from 2^{2^K} possible functions [86]) where each node has a minimum of zero to a maximum of N outputs. Such a state-space allows random network configurations which often leads to nonlinear dynamics. The total number of nodes representing the genes, here, is $N=10$, thus there are 2^{10} possible states. In RBNs, the state of the nodes in the network can be updated synchronously or asynchronously in discrete time steps. In this study, for simplicity purposes, a synchronous update rule is chosen.

RBNs have a well-defined function mapping scheme through logic (Boolean) operators which constitute the rules for connections that control the state of gene regulators. In our simulations, RBNs have different combinations of the following operators: *AND*, *OR*, *NOT, XOR, COPY*. For cBNs, operators *AND NOT, OR, NOT, COPY* are randomly placed to generate rulesets (functions). All operators have equal probability to be assigned (see Additional file 1: Text 1 and Additional file 1: Text 3 for details). Incoming edges are randomly distributed for each node with the condition that each node has at least one (thereby, connectivity is preserved) and at most two (more than one Boolean operator) edges mapped to the other nodes. Thus, outgoing edges are assigned to the nodes in a completely random fashion (allowing for the emergence of highly regulated nodes).

Although the networks that are simulated in this study carry the general characteristics of N-K models—as they are binary networks connected with Boolean functions, these networks differ in some aspects. For example, N is limited to 10, in order to generate comparable conditions with previous studies on degeneracy measures ([33, 40, 56]). The number of edges for each rule, is a randomly assigned value that can be either 1 or 2 while a node can have many outgoing edges up to N nodes. Finally, there is no constraint on the expected density of truth values (namely, bias parameter p).

Network lesioning

Degeneracy, as a strategy (or design principle [10, 21, 87]) for networks to recover their function, refers to the rearrangement of (structurally different) components in a way that function/output remains the same even after a damage. In a network with high degeneracy, there are many possible network reconfigurations that can produce/recover the function. To test the potential factors that give rise to (higher/lower) degeneracy in networks, here, we induce interventions to the systems at the network-level by lesioning the edges.

Two different lesions were introduced to the synthetic networks. In type-1 lesioning, all incoming edges from randomly chosen nodes were cut while the outgoing edges were preserved. In the second type of lesioning, all incoming and outgoing edges were cut from randomly selected nodes given the percentage of total lesioned edges. In both types of lesioning, the edges are lesioned in increments of ten percent of the total number of the nodes given a network. For example, in type-1 30% cut condition, we have lesioned all the incoming edges of the 3 randomly chosen nodes given a network of 10 nodes. Likewise, in type-2 30% cut condition, all the edges (incoming and outgoing) of 3

randomly chosen nodes (out of 10 nodes total in a network) were lesioned. By 100%-cut condition (in both lesioning types), we refer to networks where no node is connected to the other, and so the nodes are isolated.

For both lesioning types, 0%-cut condition refers to networks that are not lesioned, yet the edges are randomly disturbed (according to the method that is defined previously). This may lead some nodes to not have any edges due to random assignments of the rules, therefore, mimicking a (partially) lesioned condition.

Biologically realistic random Boolean networks: discrete to continuous

Boolean functions describe how the states of the regulators control the state of the target genes [68]. In our study, Boolean functions are randomly generated for each simulation with incremental lesioning. To execute numerical simulations, we used two types of RBNs: RBNs (classical model), and cBNs (BoolODE [84]).

Classical RBNs are discrete dynamical systems where system output consists of binary values. To induce perturbation, we set the bit-flip probability [88–92] to 0.25 for the perturbed subset. For RBNs with varying unit activity, besides the perturbation subset, all units have a bit-flip probability of 0.05 at a random time step. The total number of time steps in a simulation is 1000, and we discard the first 10 time points as a burn-in period.

The BoolODE pipeline by Pratapa et al. [68], similar to classical RBN simulations, takes input files that are randomly generated Boolean networks. Then, BoolODE systematically converts a random Boolean network into a system of SDEs that is a continuous model of gene regulation (for model specifications see Additional file 1: Text 2 and Additional file 1: Text 3). Time points in the numerical solution result in vectors of gene expression values that correspond to individual cells. That means for every analysis, each sampled time point is from a cell [68], and in this study, we sample from 990 time points (1000–10, first 10 timepoints treated as burn-in) for each gene in a single simulation and total of 1000 simulations are run for each lesioning percentage increment of 10 s (from no cut condition to all 10 genes cut) which makes 10,000 simulations for each type of lesioning and thus, 2 (lesioning type) $\times$ 10 (k subset of perturbed genes) $\times$ 11 (cut conditions, no-cut condition inclusive) $\times$ 10 synthetic cells with random gene regulatory mechanisms.

Quantification of degeneracy in neural networks

To measure degeneracy, we used the mathematical framework described by Tononi et al. [33]. In this framework, degeneracy is characterized in terms of the average mutual information between subsets of elements within a system and an output sheet (which is also a subset of network X). The output sheet is a set of randomly chosen nodes in a network and its activity is a result of the interactions the other nodes in the system. Thus, activity in the output sheet represents the behavior or the response of the whole system.

From information theory, entropy (Shannon entropy with log base 2 for binary representation) is calculated from probability density functions for subsets of X (Eq. 1). Then mutual information that measures the portion of entropy shared by the system subset X_j^k and the output O, is calculated as follows (Eq. 2):

$$H(X) = -\sum_{i=1}^{n} P(x_i) log_2 P(x_i) \tag{1}$$

$$MI\left(X_j^k; O\right) = H\left(X_j^k\right) + H(O) - H\left(X_j^k, O\right) \tag{2}$$

$H(X_j^k)$ and $H(O)$ are the entropies of X_j^k and O considered independently, whereas $H(X_j^k;O)$ is the joint entropy of subset X_j^k and output O. To measure degeneracy in the network, we need to determine the effects of the (subset of) element(s) on the entropy of the output—the behavior of the network. Since mutual information does not capture direction, however, mere calculation of mutual information is not enough to determine the contribution of the elements to the output of the system. To overcome this, perturbations (variance) are injected to the system. If no initial variance is assumed in the system, the value of mutual information between the network and the output is zero before any perturbation [33]. Variance (perturbation) is injected as uncorrelated random noise to each subset k.

Under such perturbations, mutual information of the system is computed as in *EQ3* and this procedure is repeated for all subsets of sizes $1 \leq k \leq n$. Then, degeneracy $D_N(X;O)$ of X with respect to O can be calculated as:

$$D_N(X; O) = \sum_{k=1}^{n} \left[\left\langle MI^P\left(X_j^k; O\right) \right\rangle - \left(\frac{k}{n}\right) MI^P(X; O) \right] \tag{3}$$

$MI^P(X;O)$ is MI for all elements to the output sheet, and $\langle MI^P(X_j{}^k;O)\rangle$ is the average of the contribution of each perturbed subset size k to the output sheet.

Quantification of degeneracy in cBNs

Biological RBN simulations (via BoolODE) results in continuous unit activity in terms of gene expression vectors since the simulated networks are translated into nonlinear dynamical system (Additional file 1: Text 2 and Additional file 1: Text 3, Additional file 1: Figure 1). To quantify degeneracy in RBNs we therefore discretize the gene expression vectors by taking the median activity for a unit. Activity that is above the median is set to 1, and activity below the median is set to 0.

We apply the degeneracy measures to the discretized gene expression vectors generated from the simulations. In our simulations, perturbations are systematically injected to the subset size k of genes as normally distributed (with $mean = 0$, and $standard\ deviation = 0.01$) random noise through the governing SDE. The number of elements (namely, the genes) is $n = 10$ for all simulations with output sheet consisted of the activity of $O = n/2 = 5$ elements, that is also randomized for each trial.

Partial degeneracy in random networks

Degeneracy can be measured by alternate ways that are mathematically equivalent (see other definitions in [33]). The formal definition that we use in this study requires averaging over every MI measured between each node (unit, j) which are incrementally perturbed (from $k = 1$ to $k = n$) and the output sheet for a given network structure ($\langle MI^P(X_j{}^k;O)\rangle$). However, in case where all the networks are randomly generated and the output

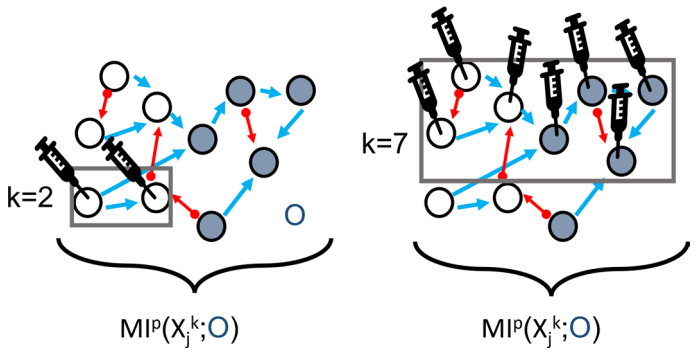

Fig. 7 Illustration of RBNs under perturbation. A network of X, composed of nodes (light and dark blue circles, n = 10) that are interconnected. Arrows represent the edges for incoming and outgoing. Light blue circles represent randomly chosen perturbation subsets of nodes for k = 2 (**a**) and k = 7 (**b**). Perturbation (represented as syringes) of the nodes in boxes with k notation, is injected as a variance (uncorrelated noise). The box with O notation represents output sheet that is also consisted of randomly chosen set of (n/2 = 5) nodes (dark blue circles). For each network, MI is calculated for perturbed set size k and the output sheet O, for all subset sizes of perturbed set noted as j

sheet units are randomly chosen, an alternate way of computing $\langle MI^P(X_j{}^k;O)\rangle$ is taking the average of MI measured for each random network that is perturbed once for a particular perturbed subset size k (see Fig. 7) in range of $1 \leq k \leq n$. This way, degeneracy is measured for a specific subset given a network rather than for all possible subset sizes. Here, we call this measurement *partial degeneracy*.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12859-022-04601-5.

Additional file 1. Supplementary Materials.

Acknowledgements
This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-20-1-0413.

Authors' contributions
All authors read and approved the final manuscript. The authors contributed equally to this work.

Funding
Not applicable.

Data availability
The links for the simulation tools and the parameters that are used in this study given in detail in Additional file 1. The datasets curated and analyzed during this work are available from the corresponding author on request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Center for Complex Systems and Brain Sciences, Florida Atlantic University, Boca Raton, FL, USA. [2]Department of Psychology, Florida Atlantic University, Boca Raton, FL, USA. [3]The Brain Institute, Florida Atlantic University, Jupiter, FL 33431, USA.

Received: 11 May 2021 Accepted: 31 January 2022
Published online: 14 February 2022

References

1. Forman-Kay JD, Mittag T. From sequence and forces to structure, function, and evolution of intrinsically disordered proteins. Structure. 2013;21(9):1492–9.
2. Sadowski MI, Jones DT. The sequence–structure relationship and protein function prediction. Curr Opin Struct Biol. 2009;19(3):357–62.
3. Pal D, Eisenberg D. Inference of protein function from protein structure. Structure. 2005;13(1):121–30.
4. Hvidsten TR, Lægreid A, Kryshtafovych A, Andersson G, Fidelis K, Komorowski J. A comprehensive analysis of the structure-function relationship in proteins based on local structure similarity. PLoS ONE. 2009;4(7):e6266.
5. Hartwell LH, Hopfield JJ, Leibler S, Murray AW. From molecular to modular cell biology. Nature. 1999;402(6761):C47-52.
6. Buzsáki G. Rhythms of the brain. Oxford: Oxford University Press; 2006.
7. Biology NRC (US) C on RO in. Molecular structure and function. opportunities in biology. Washington: National Academies Press (US); 1989.
8. Wagner A. Robustness against mutations in genetic networks of yeast. Nat Genet. 2000;24(4):355–61.
9. Subramaniam AR, Pan T, Cluzel P. Environmental perturbations lift the degeneracy of the genetic code to regulate protein levels in bacteria. Proc Natl Acad Sci. 2013;110(6):2419–24.
10. Whitacre J, Bender A. Degeneracy: a design principle for achieving robustness and evolvability. J Theor Biol. 2010;263(1):143–53.
11. Yamada T, Bork P. Evolution of biomolecular networks—lessons from metabolic and protein interactions. Nat Rev Mol Cell Biol. 2009;10(11):791–803.
12. Helsen J, Frickel J, Jelier R, Verstrepen KJ. Network hubs affect evolvability. PLoS Biol. 2019;17(1):e3000111.
13. Hintze A, Adami C. Evolution of complex modular biological networks. PLoS Comput Biol. 2008;4(2):e23.
14. Romanuk TN, Vogt RJ, Young A, Tuck C, Carscallen MW. Maintenance of positive diversity-stability relations along a gradient of environmental stress. PLoS ONE. 2010;5(4):e10378.
15. Wagner A, Wright J. Alternative routes and mutational robustness in complex regulatory networks. Biosystems. 2007;88(1):163–72.
16. Gu X. Evolution of duplicate genes versus genetic robustness against null mutations. Trends Genet. 2003;19(7):354–6.
17. Gu Z, Steinmetz LM, Gu X, Scharfe C, Davis RW, Li W-H. Role of duplicate genes in genetic robustness against null mutations. Nature. 2003;421(6918):63–6.
18. Nowak MA, Boerlijst MC, Cooke J, Smith JM. Evolution of genetic redundancy. Nature. 1997;388(6638):167–71.
19. Schneidman E, Bialek W, Berry MJ. Synergy, redundancy, and independence in population codes. J Neurosci Off J Soc Neurosci. 2003;23(37):11539–53.
20. Noppeney U, Friston KJ, Price CJ. Degenerate neuronal systems sustaining cognitive functions. J Anat. 2004;205(6):433–42.
21. Maleszka R, Mason PH, Barron AB. Epigenomics and the concept of degeneracy in biological systems. Brief Funct Genomics. 2014;13(3):191–202.
22. Kelso JAS. Multistability and metastability: understanding dynamic coordination in the brain. Philos Trans R Soc B Biol Sci. 2012;367(1591):906–18.
23. Farina SC, Kane EA, Hernandez LP. Multifunctional structures and multistructural functions: integration in the evolution of biomechanical systems. Integr Comp Biol. 2019;59(2):338–45.
24. Salathé M, Ackermann M, Bonhoeffer S. The effect of multifunctionality on the rate of evolution in yeast. Mol Biol Evol. 2006;23(4):721–2.
25. Gamfeldt L, Hillebrand H, Jonsson PR. Multiple functions increase the importance of biodiversity for overall ecosystem functioning. Ecology. 2008;89(5):1223–31.
26. Hector A, Bagchi R. Biodiversity and ecosystem multifunctionality. Nature. 2007;448(7150):188–90.
27. Atamas N, Atamas MS, Atamas F, Atamas SP. Non-local competition drives both rapid divergence and prolonged stasis in a model of speciation in populations with degenerate resource consumption. Theor Biol Med Model. 2012;9(1):56.
28. Atamas SP, Bell J. Degeneracy-driven self-structuring dynamics in selective repertoires. Bull Math Biol. 2009;71(6):1349–65.
29. Prinz AA, Bucher D, Marder E. Similar network activity from disparate circuit parameters. Nat Neurosci. 2004;7(12):1345–52.
30. Randles M, Lamb D, Odat E, Taleb-Bendiab A. Distributed redundancy and robustness in complex systems. J Comput Syst Sci. 2011;77(2):293–304.
31. Wagner A. Distributed robustness versus redundancy as causes of mutational robustness. BioEssays. 2005;27(2):176–88.
32. Hartman JL, Garvik B, Hartwell L. Principles for the buffering of genetic variation. Science. 2001;291(5506):1001–4.
33. Tononi G, Sporns O, Edelman GM. Measures of degeneracy and redundancy in biological networks. Proc Natl Acad Sci. 1999;96(6):3257–62.
34. Leonardo A. Degenerate coding in neural systems. J Comp Physiol A. 2005;191(11):995–1010.
35. Reichmann ME, Rees MW, Symons RH, Markham R. Experimental evidence for the degeneracy of the nucleotide triplet code. Nature. 1962;195:999–1000.
36. Mason PH. Degeneracy at multiple levels of complexity. Biol Theory. 2010;5(3):277–88.

37. Beverly M, Anbil S, Sengupta P. Degeneracy and neuromodulation among thermosensory neurons contribute to robust thermosensory behaviors in *Caenorhabditis elegans*. J Neurosci. 2011;31(32):11718–27.
38. Calis JJA, de Boer RJ, Keşmir C. Degenerate T-cell recognition of peptides on MHC molecules creates large holes in the T-cell repertoire. PLoS Comput Biol. 2012;8(3):e1002412.
39. Price CJ, Friston KJ. Degeneracy and cognitive anatomy. Trends Cogn Sci. 2002;6(10):416–21.
40. Man M, Zhang Y, Ma G, Friston K, Liu S. Quantification of degeneracy in Hodgkin–Huxley neurons on Newman–Watts small world network. J Theor Biol. 2016;402:62–74.
41. Marder E, Taylor AL. Multiple models to capture the variability in biological neurons and networks. Nat Neurosci. 2011;14(2):133–8.
42. Marder E, Gutierrez GJ, Nusbaum MP. Complicating connectomes: electrical coupling creates parallel pathways and degenerate circuit mechanisms. Dev Neurobiol. 2017;77(5):597–609.
43. Burley SK, Almo SC, Bonanno JB, Capel M, Chance MR, Gaasterland T, et al. Structural genomics: beyond the human genome project. Nat Genet. 1999;23(2):151–7.
44. Boeckmann B, Blatter M-C, Famiglietti L, Hinz U, Lane L, Roechert B, et al. Protein variety and functional diversity: Swiss-Prot annotation in its biological context. C R Biol. 2005;328(10):882–99.
45. Petchey OL, Gaston KJ. Functional diversity: back to basics and looking forward. Ecol Lett. 2006;9(6):741–58.
46. Edelman GM, Gally JA. Degeneracy and complexity in biological systems. Proc Natl Acad Sci. 2001;98(24):13763–8.
47. Fetz EE. Are movement parameters recognizably coded in the activity of single neurons? Behav Brain Sci. 1992;15(4):679–90.
48. Fitch HL, Kupin JJ, Kessler IJ, DeLucia J. Relating articulation and acoustics through a sinusoidal description of vocal tract shape. Speech Commun. 2003;39(3):243–68.
49. Todorov E. Optimality principles in sensorimotor control. Nat Neurosci. 2004;7(9):907–15.
50. Lee RG, van Donkelaar P. Mechanisms underlying functional recovery following stroke. Can J Neurol Sci. 1995;22(4):257–63.
51. Whitacre JM. Degeneracy: a link between evolvability, robustness and complexity in biological systems. Theor Biol Med Model. 2010;7(1):6.
52. Sporns O. Structure and function of complex brain networks. Dialogues Clin Neurosci. 2013;15(3):247–62.
53. Sporns O. Network attributes for segregation and integration in the human brain. Curr Opin Neurobiol. 2013;23(2):162–71.
54. Tononi G. Functional segregation and integration in the nervous system: theory and models. In: Franzén O, Johansson R, Terenius L, editors. Somesthesis and the neurobiology of the somatosensory cortex. Basel: Birkhäuser; 1996. p. 409–18 (**Advances in Life Sciences**).
55. Li Y, Dwivedi G, Huang W, Kemp ML, Yi Y. Quantification of degeneracy in biological systems for characterization of functional interactions between modules. J Theor Biol. 2012;302:29–38.
56. Hoel E, Levin M. Emergence of informative higher scales in biological systems: a computational toolkit for optimal prediction and control. Commun Integr Biol. 2020;13(1):108–18.
57. Shannon CE. A mathematical theory of communication. Bell Syst Tech J. 1948;27(3):379–423.
58. MacKay DJC. Information theory, inference & learning algorithms. USA: Cambridge University Press; 2002.
59. Mézard M, Montanari A. Information, physics, and computation. Information, physics, and computation. Oxford: Oxford University Press; 2009.
60. Burge C, Gifford D, Fraenkel E. Foundations of computational and systems biology [Internet]. Massachusetts Institute of Technology: MIT OpenCourseWare. 2014. https://ocw.mit.edu/courses/biology/7-91j-foundations-of-computational-and-systems-biology-spring-2014/lecture-slides/MIT7_91JS14_Lecture15.pdf.
61. Quinn CJ, Coleman TP, Kiyavash N, Hatsopoulos NG. Estimating the directed information to infer causal relationships in ensemble neural spike train recordings. J Comput Neurosci. 2011;30(1):17–44.
62. Macía J, Solé RV, Elena SF. The causes of epistasis in genetic networks. Evolution. 2012;66(2):586–96.
63. Kang S, Ma W, Li FY, Zhang Q, Niu J, Ding Y, et al. Functional redundancy instead of species redundancy determines community stability in a typical steppe of Inner Mongolia. PLoS ONE. 2015;10(12):e0145605.
64. Scherrer K, Jost J. The gene and the genon concept: a functional and information-theoretic analysis. Mol Syst Biol. 2007;3(1):87.
65. Fetzer I, Johst K, Schäwe R, Banitz T, Harms H, Chatzinotas A. The extent of functional redundancy changes as species' roles shift in different environments. Proc Natl Acad Sci. 2015;112(48):14888–93.
66. Peter IS. Methods for the experimental and computational analysis of gene regulatory networks in sea urchins. Methods Cell Biol. 2019;151:89–113.
67. Gershenson C. In: Bedau M, Husbands P, Hutton T, Kumar S, Suzuki H, editors. Workshop and tutorial proceedings. 2004. p. 160–173.
68. Pratapa A, Jalihal AP, Law JN, Bharadwaj A, Murali TM. Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data. Nat Methods. 2020;17(2):147–54.
69. Giacomantonio CE, Goodhill GJ. A Boolean model of the gene regulatory network underlying Mammalian cortical area development. PLoS Comput Biol. 2010;6(9):1000936.
70. Schaffter T, Marbach D, Floreano D. GeneNetWeaver: in silico benchmark generation and performance profiling of network inference methods. Bioinform Oxf Engl. 2011;27(16):2263–70.
71. Zalesky A, Fornito A, Bullmore ET. Network-based statistic: Identifying differences in brain networks. Neuroimage. 2010;53(4):1197–207.
72. Chai LR, Khambhati AN, Ciric R, Moore TM, Gur RC, Gur RE, et al. Evolution of brain network dynamics in neurodevelopment. Netw Neurosci. 2017;1(1):14–30.
73. De Domenico M, Sasai S, Arenas A. Mapping multiplex hubs in human functional brain networks. Front Neurosci. 2016;10:326.
74. Fornito A, Zalesky A, Breakspear M. Graph analysis of the human connectome: promise, progress, and pitfalls. Neuroimage. 2013;80:426–44.

75. Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci. 2009;10(3):186–98.
76. Medaglia JD, Lynall M-E, Bassett DS. Cognitive network neuroscience. J Cogn Neurosci. 2015;27(8):1471–91.
77. Salvador R, Anguera M, Gomar JJ, Bullmore ET, Pomarol-Clotet E. Conditional mutual information maps as descriptors of net connectivity levels in the brain. Front Neuroinform. 2010;4:115.
78. Salvador R, Martínez A, Pomarol-Clotet E, Sarró S, Suckling J, Bullmore E. Frequency based mutual information measures between clusters of brain regions in functional magnetic resonance imaging. Neuroimage. 2007;35(1):83–8.
79. Bassett DS, Bullmore ET, Meyer-Lindenberg A, Apud JA, Weinberger DR, Coppola R. Cognitive fitness of cost-efficient brain functional networks. Proc Natl Acad Sci. 2009;106(28):11747–52.
80. Moddemeijer R. On estimation of entropy and mutual information of continuous distributions. Signal Process. 1989;16(3):233–48.
81. Barabási A-L, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet. 2004;5(2):101–13.
82. Motter AE, Gulbahce N, Almaas E, Barabási A-L. Predicting synthetic rescues in metabolic networks. Mol Syst Biol. 2008;4(1):168.
83. Causey TB, Zhou S, Shanmugam KT, Ingram LO. Engineering the metabolism of Escherichia coli W3110 for the conversion of sugar to redox-neutral and oxidized products: Homoacetate production. Proc Natl Acad Sci. 2003;100(3):825–32.
84. Trinh CT, Carlson R, Wlaschin A, Srienc F. Design, construction and performance of the most efficient biomass producing *E. coli* bacterium. Metab Eng. 2006;8(6):628–38.
85. Kauffman SA. Metabolic stability and epigenesis in randomly constructed genetic nets. J Theor Biol. 1969;22(3):437–67.
86. Kauffman S. Homeostasis and differentiation in random genetic control networks. Nature. 1969;224(5215):177–8.
87. Krichmar JL. Design principles for biologically inspired cognitive robotics. Biol Inspired Cogn Archit. 2012;1:73–81.
88. Shmulevich I, Dougherty ER, Kim S, Zhang W. Probabilistic Boolean networks: a rule-based uncertainty model for gene regulatory networks. Bioinformatics. 2002;18(2):261–74.
89. Trairatphisan P, Mizera A, Pang J, Tantar AA, Schneider J, Sauter T. Recent development and biomedical applications of probabilistic Boolean networks. Cell Commun Signal CCS. 2013;11:46.
90. Peixoto TP, Drossel B. Noise in random Boolean networks. Phys Rev E. 2009;79(3):036108.
91. Zhong J, Ho DWC, Lu J, Jiao Q. Pinning controllers for activation output tracking of boolean network under one-bit perturbation. IEEE Trans Cybern. 2019;49(9):3398–408.
92. Vasic B, Ravanmehr V, Krishnan AR. An information theoretic approach to constructing robust boolean gene regulatory networks. IEEE/ACM Trans Comput Biol Bioinform. 2012;9(1):52–65.

Publisher's Note